



UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/733,541	12/11/2003	Yunbo Cao	M61.12-0556	7738
27366 7590 03/22/2007 WESTMAN CHAMPLIN (MICROSOFT CORPORATION) SUITE 1400 900 SECOND AVENUE SOUTH MINNEAPOLIS, MN 55402-3319			EXAMINER LY, CHEYNE D	
			ART UNIT 2168	PAPER NUMBER
SHORTENED STATUTORY PERIOD OF RESPONSE			MAIL DATE	DELIVERY MODE
3 MONTHS			03/22/2007	PAPER

Please find below and/or attached an Office communication concerning this application or proceeding.

If NO period for reply is specified above, the maximum statutory period will apply and will expire 6 MONTHS from the mailing date of this communication.

Office Action Summary

Application No.

10/733,541

Applicant(s)

CAO ET AL.

Examiner

Cheyne D. Ly

Art Unit

2168

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --
Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

- 1) ☒ Responsive to communication(s) filed on 13 February 2007.
- 2a) ☐ This action is **FINAL**. 2b) ☒ This action is non-final.
- 3) ☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

- 4) ☒ Claim(s) 1-3, 5-7, and 25-34 is/are pending in the application.
- 4a) Of the above claim(s) _____ is/are withdrawn from consideration.
- 5) ☐ Claim(s) _____ is/are allowed.
- 6) ☒ Claim(s) 1-3, 5-7 and 25-34 is/are rejected.
- 7) ☐ Claim(s) _____ is/are objected to.
- 8) ☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

- 9) ☐ The specification is objected to by the Examiner.
- 10) ☐ The drawing(s) filed on _____ is/are: a) ☐ accepted or b) ☐ objected to by the Examiner.
Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).
Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).
- 11) ☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

- 12) ☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).
- a) ☐ All b) ☐ Some * c) ☐ None of:
- ☐ Certified copies of the priority documents have been received.
 - ☐ Certified copies of the priority documents have been received in Application No. _____.
 - ☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).
- * See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

- | | |
|--|---|
| 1) <input type="checkbox"/> Notice of References Cited (PTO-892) | 4) <input type="checkbox"/> Interview Summary (PTO-413)
Paper No(s)/Mail Date. _____ |
| 2) <input type="checkbox"/> Notice of Draftsperson's Patent Drawing Review (PTO-948) | 5) <input type="checkbox"/> Notice of Informal Patent Application |
| 3) <input type="checkbox"/> Information Disclosure Statement(s) (PTO/SB/08)
Paper No(s)/Mail Date _____ | 6) <input type="checkbox"/> Other: _____ |

DETAILED ACTION

1. A request for continued examination under 37 CFR 1.114, including the fee set forth in 37 CFR 1.17(e), was filed in this application after final rejection. Since this application is eligible for continued examination under 37 CFR 1.114, and the fee set forth in 37 CFR 1.17(e) has been timely paid, the finality of the previous Office action has been withdrawn pursuant to 37 CFR 1.114. Applicant's submission filed on February 13, 2007 has been entered.
2. Applicants' arguments have been fully considered but they are not deemed to be persuasive. Rejections and/or objections not reiterated from previous office actions are hereby withdrawn. The following rejections and/or objections are either reiterated or newly applied. They constitute the complete set presently being applied to the instant application.
3. Claims 1-3, 5-7, and 25-34 are examined on the merit.

RESPONSE TO ARGUMENT

4. On page 6, Applicant argues via pointing to the specification on page 11, lines 17, 22, page 12, and Figures 2 and 3 have been fully considered and they have been found to be unpersuasive. It is noted that the pointed to disclosure describe the "set of related elements", however, the pointed to disclosure does not support that the "extracting" of the respective "set" from "a first string" or "a second string" respectively. The pointed to disclosure describes the extraction is directed to "a set of input strings" (page 17, lines 14-15) wherein there is not distinction of between the "a first string" or "a second string."

Art Unit: 2168

5. On pages 7-8, Applicant's arguments are directed to limitations that have been rejected for being new matter. Therefore, the prior art rejection has been maintained as directed to the claims filed, September 01, 2006. The claims, filed December 14, 2006, comprising new limitations have been fully considered. The claims, filed December 14, 2006, are free of any prior art; therefore, would be allowable if the rejections of record are overcome.

Claim Rejections - 35 USC § 112

6. The following is a quotation of the second paragraph of 35 U.S.C. 112:

The specification shall conclude with one or more claims particularly pointing out and distinctly claiming the subject matter which the applicant regards as his invention.

7. Claims 1-3, 5-7, and 25-34 are rejected under 35 U.S.C. 112, second paragraph, as being indefinite for failing to particularly point out and distinctly claim the subject matter which applicant regards as the invention.
8. Claim 1 recites the limitation "the first related set of elements" and "the second set of related elements" in line 24. There is insufficient antecedent basis for this limitation in the claim. It is noted that line 12 recites "a first set of related elements" and "a second set of related elements." The same issue is present claim 5. Claims 2, 3, 6, 7, and 25-34 are rejected for being dependent from claim 1 or 5.

CLAIM REJECTIONS - 35 USC § 101

9. 35 U.S.C. 101 reads as follows:

Whoever invents or discovers any new and useful process, machine, manufacture, or composition of matter, or any new and useful improvement thereof, may obtain a patent therefor, subject to the conditions and requirements of this title.

10. Claims 5-7 and 30-34 are rejected under 35 U.S.C. 101 because the claimed invention is directed to non-statutory algorithm type subject matter.
11. Claims 5-7 and 30-34 are nonstatutory because the claimed computer-readable medium embodying a carrier wave (page 7, line 12), wherein said carrier wave does not fall within any of the four statutory classes of invention found in 35 USC 101.
12. It is advised that Applicant amend the claims to recite "computer storage medium" as described by the specification (page 7).

Claim Rejections - 35 USC § 112

13. The following is a quotation of the first paragraph of 35 U.S.C. 112:

The specification shall contain a written description of the invention, and of the manner and process of making and using it, in such full, clear, concise, and exact terms as to enable any person skilled in the art to which it pertains, or with which it is most nearly connected, to make and use the same and shall set forth the best mode contemplated by the inventor of carrying out his invention.

14. Claims 1-3, 5-7, and 25-34 are rejected under 35 U.S.C. 112, first paragraph, as failing to comply with the written description requirement. The claim(s) contains subject matter which was not described in the specification in such a way as to reasonably convey to one skilled in the relevant art that the inventor(s), at the time the application was filed, had possession of the claimed invention. NEW MATTER.
15. Claim 1 recites the limitations of "extracting a first set of related elements..." and "extracting a second set of related elements..." wherein said limitations have not been found in the instant specification. The same issue is present claim 5.
16. In regard to claim 1, line 8, the limitation of "at least one word and a least one wildcard are positioned..." has not been found in the instant specification. It is noted claim 3, as originally filed, recites "at least one word positioned" which is distinct from the new

limitation of "at least one word and a least one wildcard." The same issue is present claim 5.

17. In regard to claim 1, lines 9-10, the limitation of "less than a specified number of words" has not been found in the instant specification. The same issue is present claim 5.

Claim Rejections - 35 USC § 103

18. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

(a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negated by the manner in which the invention was made.

19. This application currently names joint inventors. In considering patentability of the claims under 35 U.S.C. 103(a), the examiner presumes that the subject matter of the various claims was commonly owned at the time any inventions covered therein were made absent any evidence to the contrary. Applicant is advised of the obligation under 37 CFR 1.56 to point out the inventor and invention dates of each claim that was not commonly owned at the time a later invention was made in order for the examiner to consider the applicability of 35 U.S.C. 103(c) and potential 35 U.S.C. 102(e), (f) or (g) prior art under 35 U.S.C. 103(a).
20. Claims 1-3, 5-7, and 25-34 are rejected under 35 U.S.C. 103(a) as being unpatentable over Yangarber et al. (2000) (Yangarber hereafter) in view of Soderland (1999).

CLAIM INTERPRETATIONS

1. In regard to the limitations “extracting a first set of related elements...” and “extracting a second set of related elements...”, Yangarber describes the method of extracting information comprising “first element”, “second element”, “third element”, etc. and a corpus of 5,963 articles from the Wall Street Journal (page 284, column 1, last two lines, to column 2, line 22). The description of the respective elements support the respect set of elements is different from each of the related elements.

MOTIVATION TO COMBINE

2. Yangarber discloses IE systems “for finding patterns automatically from un-annotated text” (page 282, Abstract etc.). While, Soderland describes the performance of WHISK is comparable to other IE systems. For structured or semi-structure text, high or even perfect accuracy can often be achieved from a modest amount of training examples (page 236 to page 238). Therefore, one of ordinary skill in the art at the time of the invention would have been motivated by Soderland to improve the system of Yangarber for high or even perfect accuracy.

PRIOR ART

3. In regard to claim 1, Yangarber discloses a computer-implemented method of extracting information from an information source comprising a plurality of documents (page 283, column 2, lines 5-9, especially, “documents”, and page 286, column 2, section 4.2, especially, “corpus of 100 documents”);

Accessing strings of text in he information source (page 283, column 2, lines 10-17, especially, “patterns”);

Art Unit: 2168

Comparing the strings of text in the information source with generalized extraction patterns and identifying a plurality of strings in the information source that match at least one generalized extraction pattern (page 283, column 1, lines 1-5, especially, “generalize them into patterns”, and column 2, lines 32-34, especially, “Generalize each pattern...”), the generalized extraction patterns including words and wildcards (page 285, column 1, section 3.5, especially, “wild-cards”);

Extracting a first set of related elements of text pertaining to a topic from a first string of the plurality of strings based on a corresponding set of related elements pertaining to the topic in the at least one generalized extraction pattern, the first string being associated with a first document in the plurality of documents (page 282, column 1, last 5 lines, and page 285, column 1, section 3.4, especially, “management succession” as “topic”);

Extracting a second set of related elements of text pertaining to the topic from a second string of plurality of strings based on the corresponding set of related elements in the at least one generalized extraction pattern, the second string being associated with a second document in the plurality of documents, wherein at least one of the related elements of text in the first set of related elements is different from each of the related elements of text in the second set of related elements of text (page 285, column 1, section 3.5, to column 2, last line,); and

Outputting the first related set of elements and the second set of related elements (page 285, column 1, last 3 lines).

4. However, Yangarber does not explicitly describe that the wildcards denote that at least one word in an individual string can be skipped in order to match the individual string to an individual generalized extraction pattern. Soderland discloses he wildcards denote that at least one word in an individual string can be skipped in order to match the individual string to an individual generalized extraction pattern (page 238, section 2.1, especially, “wildcard ‘*’ means to skip any number of characters...”). Therefore, it would have been obvious to one of ordinary skill in the art at the time of the invention to make and use the system of Yangarber with the wildcard feature described by Soderland for high or even perfect accuracy.

5. In regard to claim 2, Yangarber describes processing the first related set of elements and the second set of related elements to analyze data in the information source (page 284, column 1, last two lines, to column 2, line 22). Therefore, it would have been obvious to one of ordinary skill in the art at the time of the invention to make and use the system of Yangarber with the wildcard feature described by Soderland for high or even perfect accuracy.

6. In regard to claim 3, Yangarber describes all the limitations of said claims except for the limitation of “the generalized extraction patterns there is at least one word positioned between at least one of the elements and the indication.” Soderland describes the limitation of “the generalized extraction patterns there is at least one word positioned between at least one of the elements and the indication” (page 238, section 2.1, to page 241, especially, Figures 6, 8, and 10), as required by the claimed invention. Therefore, it would have been obvious to one of ordinary skill in the art at the time of the invention to make and use the

system of Yangarber with the wildcard feature described by Soderland for high or even perfect accuracy.

7. In regard to claims 5-7, Yangarber and Soderland render the claimed invention obvious in light of the above cited prior art as a whole. Further, Soderland specifically describes it is well known that IE systems such as WHISK comprises the claimed “computer readable medium” (page 261, section 6.1, especially, “database applications and software agent technology can be enabled by IE systems”). Therefore, it would have been obvious to one of ordinary skill in the art at the time of the invention to make and use the system of Yangarber with the wildcard feature described by Soderland for high or even perfect accuracy.

8. In regard to claim 25, Yangarber describes each of the elements of the first set of related elements of text are different from each of the elements of the second set of related elements of text (page 284, column 1, last two lines, to column 2, line 22). Therefore, it would have been obvious to one of ordinary skill in the art at the time of the invention to make and use the system of Yangarber with the wildcard feature described by Soderland for high or even perfect accuracy.

9. In regard to claim 26, Yangarber describes the corresponding related set of elements refer to general elements pertaining to the topic and the first set of related elements and the second set of related elements refer to specific text associated with the general elements (page 284, column 1, last two lines, to page 285, column 1, section 3.4, especially the describe elements, and generalization and concept classes). Therefore, it would have been obvious to one of ordinary skill in the art at the time of the invention to make and use the system of Yangarber with the wildcard feature described by Soderland for high or even perfect accuracy.

Art Unit: 2168

10. In regard to claims 27 and 28, Yangarber describes all the limitations of said claims except for the limitation of “company/product pair.” Soderland describes the limitation of “company/product pair” (page 235 in its entirety, especially, “Z Corp” and “maker of semiconductors”) as required by the claimed invention. Therefore, it would have been obvious to one of ordinary skill in the art at the time of the invention to make and use the system of Yangarber with the wildcard feature described by Soderland for high or even perfect accuracy.

11. In regard to claim 29, Yangarber describes “the plurality of documents include at least one of a collection of documents, news articles...” (page 284, column 1, last two lines, to column 2, line 22, especially, the Wall Street Journal). Therefore, it would have been obvious to one of ordinary skill in the art at the time of the invention to make and use the system of Yangarber with the wildcard feature described by Soderland for high or even perfect accuracy.

12. In regard to claims 30-34 Yangarber and Soderland render the claimed in invention obvious in light of the above cited prior art as a whole. Further, Soderland specifically describes it is well known that IE systems such as WHISK comprises the claimed “computer readable medium” (page 261, section 6.1, especially, “database applications and software agent technology can be enabled by IE systems”). Therefore, it would have been obvious to one of ordinary skill in the art at the time of the invention to make and use the system of Yangarber with the wildcard feature described by Soderland for high or even perfect accuracy.

Art Unit: 2168

CONCLUSION

13. Patent applicants with problems or questions regarding electronic images that can be viewed in the Patent Application Information Retrieval system (PAIR) can now contact the USPTO's Patent Electronic Business Center (Patent EBC) for assistance. Representatives are available to answer your questions daily from 6 am to midnight (EST). The toll free number is (866) 217-9197. When calling please have your application serial or patent number, the type of document you are having an image problem with, the number of pages and the specific nature of the problem. The Patent Electronic Business Center will notify applicants of the resolution of the problem within 5-7 business days. Applicants can also check PAIR to confirm that the problem has been corrected. The USPTO's Patent Electronic Business Center is a complete service center supporting all patent business on the Internet. The USPTO's PAIR system provides Internet-based access to patent application status and history information. It also enables applicants to view the scanned images of their own application file folder(s) as well as general patent information available to the public.

14. For all other customer support, please call the USPTO Call Center (UCC) at 800-786-9199. The USPTO's official fax number is 571-272-8300.

15. Any inquiry concerning this communication or earlier communications from the examiner should be directed to C. Dune Ly, whose telephone number is (571) 272-0716. The examiner can normally be reached on Monday-Friday from 8 A.M. to 4 P.M.

16. If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Tim Vo, can be reached on (571) 272-3642.

C. Dune Ly
Patent Examiner



Application/Control Number: 10/733,541

Page 12

Art Unit: 2168

3/13/07